
WIREH

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Dec 23 10:13:48 1999; MasPar time 4.30 Seconds

Tabular output not generated. 76.214 Million cell updates/sec

Title: >US-09-177-843-2

Description: (1-6) from US09177843.pep

Perfect Score: 41

Sequence: 1 GRGESP 6

Scoring table: PAM 150
Gap 15

Searched: 179066 seqs, 54579741 residues

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: sprembl9

1:sp-archaea 2:sp-bacteria 3:sp-fungi 4:sp-human
5:sp-invertebrate 6:sp-mammal 7:sp-mhc 8:sp-organelle
9:sp-phage 10:sp-plant 11:sp-rodent 12:sp-unclassified
13:sp-vertebrate 14:sp-virus

Statistics: Mean 18.265; Variance 16.691; scale 1.094

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description | Pred. No. |
|------------|-------|-------------|--------|-----------|------------------------|-----------|
| 1 | 41 | 100.0 | 95 | 1 Q50799 | COBALT TRANSPORT PROTE | 1.96e+00 |
| 2 | 41 | 100.0 | 347 | 4 Q99703 | ATAXIN-2 RELATED PROTE | 1.96e+00 |
| 3 | 41 | 100.0 | 625 | 11 Q99053 | FILAGGRIN (FRAGMENT) | 1.96e+00 |
| 4 | 40 | 97.6 | 1095 | 4 Q99458 | NOTCH4 (FRAGMENT) | 3.73e+00 |
| 5 | 40 | 97.6 | 2478 | 13 Q93406 | FIBONECTIN | 3.73e+00 |
| 6 | 39 | 95.1 | 316 | 2 O83343 | RIBOSE/GALACTOSE ABC T | 7.01e+00 |
| 7 | 39 | 95.1 | 469 | 2 Q58436 | S-LAYER-LIKE PROTEIN | 7.01e+00 |
| 8 | 38 | 92.7 | 123 | 5 Q18672 | HYPOTHETICAL PROTEIN C | 1.30e+01 |
| 9 | 38 | 92.7 | 250 | 11 Q61571 | FILAGGRIN (FRAGMENT) | 1.30e+01 |
| 10 | 38 | 92.7 | 255 | 11 Q61570 | FILAGGRIN (FRAGMENT) | 1.30e+01 |
| 11 | 38 | 92.7 | 339 | 5 Q18218 | CODED FOR BY C. ELEGAN | 1.30e+01 |
| 12 | 38 | 92.7 | 356 | 7 Q30227 | MHC CLASS I AOTR-G*02 | 1.30e+01 |
| 13 | 38 | 92.7 | 511 | 14 Q91332 | NUCLEAR ANTIGEN EBNA-1 | 1.30e+01 |
| 14 | 38 | 92.7 | 699 | 10 Q64525 | KUP8H12R.12 PROTEIN | 1.30e+01 |
| 15 | 38 | 92.7 | 1095 | 4 Q60300 | KIAA0553 PROTEIN (FRAG | 1.30e+01 |
| 16 | 38 | 92.7 | 1190 | 11 P97573 | INOSITOL POLYPHOSPHATE | 1.30e+01 |
| 17 | 38 | 92.7 | 1191 | 11 Q61181 | INOSITOL POLYPHOSPHATE | 1.30e+01 |
| 18 | 38 | 92.7 | 1307 | 13 Q57683 | 146KDA NUCLEAR PROTEIN | 1.30e+01 |
| 19 | 38 | 92.7 | 3766 | 5 Q17551 | SIMILAR TO REGULATOR O | 1.30e+01 |
| 20 | 37 | 90.2 | 235 | 1 P71164 | ASPARTATE RACEMASE | 2.39e+01 |

| | | | | | | |
|----|----|------|------|-----------|------------------------|----------|
| 21 | 37 | 90.2 | 546 | 6 Q29111 | SCAVENGER-RECEPTOR PRO | 2.39e+01 |
| 22 | 37 | 90.2 | 568 | 2 Q44106 | MAJOR SURFACE PROTEIN | 2.39e+01 |
| 23 | 37 | 90.2 | 682 | 5 Q22537 | SIMILARITY TO COLLAGEN | 2.39e+01 |
| 24 | 37 | 90.2 | 804 | 6 Q29113 | SCAVENGER-RECEPTOR PRO | 2.39e+01 |
| 25 | 37 | 90.2 | 1272 | 13 Q29024 | NEUROFASCIN PRECURSOR | 2.39e+01 |
| 26 | 37 | 90.2 | 1280 | 13 Q30933 | NEURON-GLIA CELL ADHES | 2.39e+01 |
| 27 | 37 | 90.2 | 1369 | 13 Q42414 | NEUROFASCIN PRECURSOR | 2.39e+01 |
| 28 | 36 | 87.8 | 181 | 1 Q27963 | CONSERVED HYPOTHETICAL | 4.34e+01 |
| 29 | 36 | 87.8 | 253 | 5 Q95005 | C36B1.4 PROTEIN | 4.34e+01 |
| 30 | 36 | 87.8 | 551 | 10 Q81469 | T15F16.13 PROTEIN | 4.34e+01 |
| 31 | 36 | 87.8 | 633 | 2 Q85869 | GROUP11 INTRON-ASSOCIA | 4.34e+01 |
| 32 | 36 | 87.8 | 814 | 2 Q86591 | IRON-UTAKE FACTOR | 4.34e+01 |
| 33 | 36 | 87.8 | 2752 | 11 Q63461 | PROLINE-RICH PROTEIN P | 4.34e+01 |
| 34 | 36 | 87.8 | 5027 | 11 Q63460 | PROLINE-RICH PROTEIN (| 4.34e+01 |
| 35 | 35 | 85.4 | 58 | 4 Q14489 | RIBOSOMAL PROTEIN S10 | 7.77e+01 |
| 36 | 35 | 85.4 | 95 | 5 Q24411 | MYOSIN HEAVY CHAIN 1 (| 7.77e+01 |
| 37 | 35 | 85.4 | 222 | 10 Q80340 | ETHYLENE RESPONSIVE EL | 7.77e+01 |
| 38 | 35 | 85.4 | 441 | 4 Q75685 | GNAS1 PROTEIN (FRAGEN | 7.77e+01 |
| 39 | 35 | 85.4 | 501 | 14 Q98111 | ULI[-1] PROTEIN | 7.77e+01 |
| 40 | 35 | 85.4 | 616 | 1 Q26384 | ATP-DEPENDENT CLP PROT | 7.77e+01 |
| 41 | 35 | 85.4 | 711 | 2 Q96855 | HYPOTHETICAL 74.5 KD P | 7.77e+01 |
| 42 | 35 | 85.4 | 839 | 5 Q26024 | TOPOISOMERASE I | 7.77e+01 |
| 43 | 35 | 85.4 | 904 | 10 Q48541 | LIMIT DEXTRINASE | 7.77e+01 |
| 44 | 35 | 85.4 | 1014 | 5 Q26152 | V-SERA 2 | 7.77e+01 |
| 45 | 35 | 85.4 | 1201 | 5 Q18392 | MUSCLE MYOSIN HEAVY CH | 7.77e+01 |

ALIGNMENTS

| RESULT | 1 | PRELIMINARY; | PRT; | 95 AA. |
|--|---|--------------|------|---------|
| ID | Q50799 | | | |
| AC | Q50799 | | | |
| DT | 01-JUN-1998 (TREMBLREL. 06, CREATED) | | | |
| DT | 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE) | | | |
| DT | 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE) | | | |
| DE | COBALT TRANSPORT PROTEIN CBIN. | | | |
| GN | CBIN. | | | |
| OS | METHANOBACTERIUM THERMOAUTOTROPHICUM (STRAIN MARBURG) | | | |
| OC | ARCHAEA; EURYARCHAEOTA; METHANOBACTERIALES; METHANOBACTERIACEAE; | | | |
| RC | METHANOBACTERIUM. | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RA | EBERHARDT S., KORN S., LOTTSPEICH F., BACHER A.; | | | |
| RL | J.BACTERIOL. 179:2938-2943(1997). | | | |
| CC | !- FUNCTION: MAY BE INVOLVED WITH COBALT TRANSPORT IN ASSOCIATION | | | |
| CC | WITH COBALAMIN BIOSYNTHESIS. | | | |
| CC | !- PATHWAY: COBALAMIN BIOSYNTHESIS. | | | |
| CC | !- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. | | | |
| DR | EMBL; X94292; E237183; - | | | |
| KW | COBALAMIN BIOSYNTHESIS; TRANSPORT; COBALT TRANSPORT; TRANSMEMBRANE. | | | |
| FT | TRANSMEM 7 27 POTENTIAL. | | | |
| FT | TRANSMEM 67 87 POTENTIAL. | | | |
| SQ | SEQUENCE 95 AA; 10254 MW; 81CE9869 CRC32; | | | |
| Query Match 100.0%; Score 41; DB 1; Length 95; | | | | |
| Best Local Similarity 100.0%; Pred. No. 1.96e+00; | | | | |
| Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | | | | |
| Db | 89 GRGESP 94 | | | |
| | | | | |
| Qy | 1 GRGESP 6 | | | |
| RESULT 2 | | | | |
| ID | Q99703 | PRELIMINARY; | PRT; | 347 AA. |
| AC | Q99703 | | | |
| DT | 01-MAY-1997 (TREMBLREL. 03, CREATED) | | | |
| DT | 01-MAY-1997 (TREMBLREL. 03, LAST SEQUENCE UPDATE) | | | |
| DT | 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE) | | | |
| DE | ATAXIN-2 RELATED PROTEIN (FRAGMENT) | | | |
| OS | HOMO SAPIENS (HUMAN) | | | |
| OC | EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES; | | | |
| OC | CATARRHINI; HOMINIDAE; HOMO. | | | |

```

RN SEQUENCE FROM N.A.
RP MEDLINE; 97051920.
RA PULST S.M., NECHIPORUK A., NECHIPORUK T., GISPET S., CHEN X.N.,
RA LOPES-CENDES I., PEARLMAN S., STARKMAN S., OROZCO-DIAZ G., LUNKES A.,
RA DEJONG P., ROULEAU G.A., AUBURGER G., KORENBERG J.R., FIGUEROA C.,
RA SAHBA S.;
RT "Moderate expansion of a normally biallelic trinucleotide repeat in
RT spinocerebellar ataxia type 2.";
RL NAT. GENET. 14:269-276(1996).
DR EMBL; U70671; G1679686; -.
FT NON_TER 1
SQ SEQUENCE 347 AA; 37805 MW; 9BD963F2 CRC32;

Query Match 100.0%; Score 41; DB 4; Length 347;
Best Local Similarity 100.0%; Pred. No. 1.96e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 283 GRGESP 288
|||||
QY 1 GRGESP 6

RESULT 3
ID Q99053 PRELIMINARY; PRT; 625 AA.
AC Q99053;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE FLAGGRIN (FRAGMENT)
OS RATTUS NORVEGICUS (RAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;
OC SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.
[1]
RN SEQUENCE FROM N.A.
RP TISSUE-LIVER;
RX MEDLINE; 90274870.
RA HAYDOCK P.V., DALE B.A.;
RT "Flaggrin, an intermediate filament-associated protein: structural
RT and functional implications from the sequence of a cDNA from rat.";
RL DNA CELL BIOL. 9:251-261(1990).
CC -1- FUNCTION: FLAGGRIN AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND
CC PROMOTES DISULFID-BOND FORMATION AMONGST THE INTERMEDIATE
CC FILAMENTS DURING TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.
CC -1- FLAGGRIN IS INITIALLY SYNTHESIZED AS A LARGE, INSOLUBLE, HIGHLY
CC PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM COPIES OF 380 AA,
CC WHICH ARE SEPARATED BY "LINKERS" OF 26 AA.
CC THE PRECURSOR IS DEPOSITED AS KERATOHVALIN GRANULES.
CC DURING TERMINAL DIFFERENTIATION IT IS DEPHOSPHORYLATED &
CC PROTEOLYTICALLY CLEAVED.
CC -1- TISSUE SPECIFICITY: EPIDERMIS, PALATE AND FORESTOMACH.
DR EMBL; M21759; G204144; -.
DR PIR; A34615; A34615.
KW DEVELOPMENTAL PROTEIN; PHOSPHORYLATION; POLYPROTEIN.
FT NON_TER 1
SQ SEQUENCE 625 AA; 65957 MW; 6CD02B43 CRC32;

Query Match 100.0%; Score 41; DB 11; Length 625;
Best Local Similarity 100.0%; Pred. No. 1.96e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 126 GRGESP 131
|||||
QY 1 GRGESP 6

RESULT 4
ID Q99458 PRELIMINARY; PRT; 1095 AA.
AC Q99458;
DT 01-MAY-1997 (TREMBLREL. 03, CREATED)
DT 01-MAY-1997 (TREMBLREL. 03, LAST SEQUENCE UPDATE)
DT 01-JAN-1999 (TREMBLREL. 09, LAST ANNOTATION UPDATE)
DE NOTCH4 (FRAGMENT).

```

```

GN NOTCH4.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;
OC CATARRHINI; HOMINIDAE; HOMO.
[1]
RN SEQUENCE FROM N.A.
RP TISSUE=PLACENTA;
RX MEDLINE; 97311416.
RA SUGAYA K., SASANUMA S., NOHATA J., KIMURA T., FUKAGAWA T.,
RA NAKAMURA Y., ANDO A., INOKO H., IKEMURA T., MITA K.;
RT "Gene organization of human NOTCH4 and (CTG)n polymorphism in this
RT human counterpart gene of mouse proto-oncogene Int3.";
RL GENE 189:235-244(1997).
DR EMBL; D63395; D1010354; -.
DR PROSITE; PS01186; EGF_2; 5.
DR PFAM; PF00008; EGF; 5.
DR PFAM; PF00023; ank; 5.
DR PFAM; PF00066; notch; 2.
KW GLYCOPROTEIN.
FT NON_TER 1
SQ SEQUENCE 1095 AA; 114807 MW; 598EA191 CRC32;

Query Match 97.68; Score 40; DB 4; Length 1095;
Best Local Similarity 83.3%; Pred. No. 3.73e+00;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 547 GRGESP 552
|||||
QY 1 GRGESP 6

RESULT 5
ID Q93406 PRELIMINARY; PRT; 2478 AA.
AC Q93406;
DT 01-NOV-1998 (TREMBLREL. 08, CREATED)
DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE FIBRONECTIN.
GN ZFN.
OS BRACHYDANIO RERIO (ZEBRAFISH) (ZEBRA DANIO).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ACTINOPTERYGII; NEOPTERYGII;
OC TELEOSTEI; EUTELEOSTEI; OSTARIOPHYSI; CYPRINIFORMES; CYPRINOIDEA;
OC CYPRINIDAE; RASBORINAE; DANIO.
[1]
RN SEQUENCE FROM N.A.
RA ZHAO Q., COLLODI P.;
RT "Characterization and expression of zebrafish fibronectin.";
RL SUBMITTED (JUL-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; AF081128; G3420846; -.
DR PROSITE; PS00023; FIBRONECTIN_2; 2.
SQ SEQUENCE 2478 AA; 271652 MW; 1991DF70 CRC32;

Query Match 97.68; Score 40; DB 13; Length 2478;
Best Local Similarity 83.3%; Pred. No. 3.73e+00;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 1613 GRGESP 1618
|||||
QY 1 GRGESP 6

RESULT 6
ID Q83343 PRELIMINARY; PRT; 316 AA.
AC Q83343;
DT 01-NOV-1998 (TREMBLREL. 08, CREATED)
DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE RIBOSE/GALACTOSE ABC TRANSPORTER, PERMEASE PROTEIN (RBSC-2).
GN TP0323.
OS TREPONEMA PALLIDUM.
OC BACTERIA; SPIROCHAETALES; SPIROCHAETACEAE; TREPONEMA.
[1]
RN SEQUENCE FROM N.A.

```

RX MEDLINE: 98332770.
 RA FRASER C.M., NORRIS S.J., WEINSTOCK G.M., WHITE O., SUTTON G.G.,
 RA DODSON R., GWINN M., HICKEY E.K., CLAYTON R., KETCHUM K.A.,
 RA SODERGREN E., HARDHAM J.M., MCLEOD M.P., SALZBERG S., PETERSON J.,
 RA KHALAK H., RICHARDSON D., HOWELL J.K., CHIDAMBARAM M., UTERBACK T.,
 RA MCDONALD L., ARTTACH P., BOWMAN C., COTTON M.D., FUJII C., GARLAND S.,
 RA HATCH B., HORST K., ROBERTS K., WATTHEY L., WEIDMAN J., SMITH H.O.,
 RA VENTER J.C.;
 RT "Complete Genome Sequence of Treponema pallidum, the Syphilis
 Spirochete";
 RL SCIENCE 281:375-388(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA FRASER C.M., NORRIS S.J., WEINSTOCK G.M., WHITE O., SUTTON G.G.,
 RA DODSON R., GWINN M., HICKEY E.K., CLAYTON R., KETCHUM K.A.,
 RA SODERGREN E., HARDHAM J.M., MCLEOD M.P., SALZBERG S., PETERSON J.,
 RA KHALAK H., RICHARDSON D., HOWELL J.K., CHIDAMBARAM M., UTERBACK T.,
 RA MCDONALD L., ARTTACH P., BOWMAN C., COTTON M.D., FUJII C., GARLAND S.,
 RA HATCH B., HORST K., ROBERTS K., WATTHEY L., WEIDMAN J., SMITH H.O.,
 RA VENTER J.C.;
 RL SUBMITTED (MAR-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL; AE001212; G3322600; -.
 SQ SEQUENCE 316 AA; 33824 MW; CE7D88A0 CRC32;

Query Match 95.1%; Score 39; DB 2; Length 316;
 Best Local Similarity 83.3%; Pred. No. 7.01e+00;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 298 GRGEAP 303
 ||||:1
 QY 1 GRGESP 6

RESULT 7
 ID Q56436 PRELIMINARY; PRT; 469 AA.
 AC Q56436;
 DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
 DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
 DT 01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
 DE S-LAYER-LIKE PROTEIN.
 GN SLPN.
 OS THERMUS AQUATICUS (SUBSP. THERMOPHILUS).
 OC BACTERIA; THERMUS/DINOCOCCUS GROUP; THERMUS GROUP; THERMUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-HB8;
 RA OLABARRIA G., FERNANDEZ-HERRERO L.A., CARRASCOSA J.L., BERENGUER J.;
 RL J. BACTERIOL. 178:357-365(1995).
 DR EMBL; X90369; G993026; -.
 KW S-LAYER.
 SQ SEQUENCE 469 AA; 52131 MW; 87F62633 CRC32;

Query Match 95.1%; Score 39; DB 2; Length 469;
 Best Local Similarity 83.3%; Pred. No. 7.01e+00;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 178 GRGEAP 183
 ||||:1
 QY 1 GRGESP 6

RESULT 8
 ID Q18672 PRELIMINARY; PRT; 123 AA.
 AC Q18672;
 DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
 DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE HYPOTHETICAL PROTEIN C47D12.10.
 GN C47D12.10.
 OS CAENORHABDITIS ELEGANS.
 OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;
 OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.
 RN [1]

RP SEQUENCE FROM N.A.
 RA GAJADSTY S.;
 RL SUBMITTED (MAR-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA MEDLINE: 94150718.
 RA WILSON R., AINSOUGH R., ANDERSON K., BAYNES C., BERKS M.,
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
 RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
 RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 elegans";
 RL NATURE 368:32-38(1994).
 DR EMBL; Z69902; E348316; -.
 KW HYPOTHETICAL PROTEIN.
 SQ SEQUENCE 123 AA; 14093 MW; D88588D7 CRC32;

Query Match 92.7%; Score 38; DB 5; Length 123;
 Best Local Similarity 83.3%; Pred. No. 1.30e+01;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 67 GRGOSP 72
 ||||:11
 QY 1 GRGESP 6

RESULT 9
 ID Q61571 PRELIMINARY; PRT; 250 AA.
 AC Q61571;
 DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
 DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE FILAGGRIN (FRAGMENT).
 OS MUS MUSCULUS (MOUSE).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;
 OC SCIUROGNATHI; MURIDAE; MURINAE; MUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 90130423.
 RA ROTHNAGEL J.A., STEINERT P.M.;
 RT "The structure of the gene for mouse filaggrin and a comparison of
 the repeating units";
 RL J. BIOL. CHEM. 265:1862-1865(1990).
 DR EMBL; M32301; G193312; -.
 FT NON_TER 1 1
 FT NON_TER 250 250
 SQ SEQUENCE 250 AA; 26479 MW; 0664963A CRC32;

Query Match 92.7%; Score 38; DB 11; Length 250;
 Best Local Similarity 83.3%; Pred. No. 1.30e+01;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 186 GRGOSP 191
 ||||:11
 QY 1 GRGESP 6

RESULT 10
 ID Q61570 PRELIMINARY; PRT; 255 AA.
 AC Q61570;
 DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
 DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE FILAGGRIN (FRAGMENT).
 GN FLG.
 OS MUS MUSCULUS (MOUSE).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;

```
OC SCIURGNATHI; MURIDAE; MURINAE; MUS.
RN
RP SEQUENCE FROM N.A.
RX MEDLINE; 90130423.
RA ROTHNAGEL J.A., STEINERT P.M.;
RT "The structure of the gene for mouse filaggrin and a comparison of
the repeating units";
RL J. BIOL. CHEM. 265:1862-1865(1990).
DR EMBL; M32300; G193310; -.
DR MGD; MGI:95553; FIG.
FT NON_TER 1
FT NON_TER 255
FT SEQUENCE 255 AA; 27162 MW; AB28D802 CRC32;
SQ
Query Match 92.7%; Score 38; DB 11; Length 255;
Best Local Similarity 83.3%; Pred. No. 1.30e+01;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Db 191 GRGOSP 196
QY 1 GRGESP 6
RESULT 11
ID Q18218 PRELIMINARY; PRT; 339 AA.
AC Q18218;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE CODED FOR BY C. ELEGANS CDNA CEMSH02F.
GN C26E5.11.
OS CAENORHABDITIS ELEGANS.
OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;
OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.
RN
RP SEQUENCE FROM N.A.
RX STRAIN-BRISTOL N2;
RA WILSON R., AINSOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA JONES M., KERSHAW J., KRISTEN J., LAISTER N., LATREILLE P.,
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA SWALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
RA THIDERY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN-BRISTOL N2;
RA FULTON L.;
RL SUBMITTED (NOV-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN-BRISTOL N2;
RA WATERSTON R.;
RL SUBMITTED (AUG-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; U13875; G532812; -.
SQ SEQUENCE 339 AA; 38108 MW; 6F6B4FF0 CRC32;
Query Match 92.7%; Score 38; DB 5; Length 339;
Best Local Similarity 83.3%; Pred. No. 1.30e+01;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Db 67 GRGOSP 72
QY 1 GRGESP 6
RESULT 12
ID Q30227 PRELIMINARY; PRT; 356 AA.
AC Q30227;
```

```
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE MHC CLASS I AOTR-G*02 (FRAGMENT).
GN AOTR-G.
OS AOTUS TRIVIRGATUS (NIGHT MONKEY) (DOUROUCOULI).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;
OC PLATYRRHINI; CEBIDAE; AOTINAE; AOTUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 98070787.
RA CADAVID L.F., SHUFFLEBOTHAM C., RUIZ F.J., YEAGER M., HUGHES A.L.,
RA WATKINS D.I.;
RT "Evolutionary instability of the major histocompatibility complex
class I loci in New World primates.";
RL PROC. NATL. ACAD. SCI. U.S.A. 94:14536-14541(1997).
DR EMBL; U59645; G1389917; -.
DR PFAM; PF00047; Ig; 1.
DR PFAM; PF00129; MHC_I; 1.
DR MHC.
FT NON_TER 1
FT SEQUENCE 356 AA; 39726 MW; B3A2E8B0 CRC32;
SQ
Query Match 92.7%; Score 38; DB 7; Length 356;
Best Local Similarity 83.3%; Pred. No. 1.30e+01;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Db 32 GRGESP 37
QY 1 GRGESP 6
RESULT 13
ID O91332 PRELIMINARY; PRT; 511 AA.
AC O91332;
DT 01-NOV-1998 (TREMBLREL. 08, CREATED)
DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE NUCLEAR ANTIGEN EBNA-1.
OC CERCOPITHECINE HERPESVIRUS 15 (RHEBUS EPSTEIN BARR VIRUS).
OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; HERPESVIRIDAE;
OC GAMMAHERPESVIRINAE; LYMPHOCRYPTOVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-LCL8664;
RA MOGHADDAM A., ANNIS B., WANG F.;
RT "Rhesus lymphocryptovirus homolog of Epstein-Barr virus EBNA-1.";
RL SUBMITTED (MAR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; U93909; G3342234; -.
SQ SEQUENCE 511 AA; 49827 MW; 343329C1 CRC32;
Query Match 92.7%; Score 38; DB 14; Length 511;
Best Local Similarity 83.3%; Pred. No. 1.30e+01;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Db 212 GRGESP 217
QY 1 GRGESP 6
RESULT 14
ID O64525 PRELIMINARY; PRT; 699 AA.
AC O64525;
DT 01-AUG-1998 (TREMBLREL. 07, CREATED)
DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE YUP8H12R.12 PROTEIN.
GN YUP8H12R.12.
OS ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).
OC EUKARYOTA; VIRIDIPANTAE; STREPTOPHYTA; EMBRYOPHYTA; TRACHEOPHYTA;
OC EUPHYLLOPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; EUDICOTYLEDONS; ROSIDAE;
OC CAPPARALES; BRASSICACEAE; ARABIDOPSIS.
RN [1]
```

Search completed: Thu Dec 23 10:14:27 1999
Job time : 39 secs.

```

RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA THEOLOGIS A.; VYSOTSKATA V.S., OSBORNE B.I., SCHWARTZ J.R.,
RA FEDERSPIEL N.A., KWAN A., TORIUMI M., YU G., CUI, O, ARAUJO R.,
RA CHUNG E., DEWER K., DIETRICH F., ECKER J.R., MARZIALI A., OEFNER P.,
RA DAVIS R.W.;
RL SUBMITTED (OCT-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA THEOLOGIS A.;
RL SUBMITTED (SEP-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA THEOLOGIS A.;
RL SUBMITTED (OCT-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA THEOLOGIS A.;
RL SUBMITTED (OCT-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA THEOLOGIS;
RL SUBMITTED (MAY-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; AC002986; G3152593; -
SQ SEQUENCE 699 AA; 76929 MW; 10513A24 CRC32;

Query Match          92.7%; Score 38; DB 10; Length 699;
Best Local Similarity 83.3%; Pred. No. 1.30e+01;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 691 GRGSP 696
   |||:|
Qy 1 GRGSP 6

RESULT 15
ID O60300 PRELIMINARY; PRT; 1095 AA.
AC O60300;
DT 01-AUG-1998 (TREMREL. 07, CREATED)
DT 01-AUG-1998 (TREMREL. 07, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMREL. 08, LAST ANNOTATION UPDATE)
DE KIAA0553 PROTEIN (FRAGMENT).
GN KIAA0553.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;
OC CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE; 98290545.
RA NAGASE T., ISHIRAWA K., MIYAJIMA N., TANAKA A., KOTANI H., NOMURA N.,
RA OHARA O.;
RT "Prediction of the coding sequences of unidentified human genes. IX.
RT The complete sequences of 100 new cDNA clones from brain which can
RT code for large proteins in vitro.";
RL DNA RES. 5:31-39(1998).
DR EMBL; AB011125; D1026409; -
FT NON_TER
SQ SEQUENCE 1095 AA; 118835 MW; 43474254 CRC32;

Query Match          92.7%; Score 38; DB 4; Length 1095;
Best Local Similarity 83.3%; Pred. No. 1.30e+01;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 635 GRGSP 640
   |||:|
Qy 1 GRGSP 6

```

THIS PAGE BLANK (USPTO)